Exploring the metabolic potential of Penicillium rubens

Viggiano, Annarita

DOI:
10.33612/diss.126598491

CHAPTER 2

Pathway for the biosynthesis of the pigment chrysogine by *Penicillium rubens*

Annarita Viggiano[1], Oleksandr Salo[1], Hazrat Ali[1]*, Wiktor Szymanski[2,3], Peter P. Lankhorst[4], Yvonne Nygård[1]*, Roel A. L. Bovenberg[4,5], Arnold J. M. Driessen[1]

[1]Molecular Microbiology, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, Groningen, The Netherlands
[2]Department of Radiology, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands
[3]Centre for Systems Chemistry, Stratingh Institute for Chemistry, University of Groningen, Groningen, The Netherlands
[4]DSM Biotechnology Centre, Delft, The Netherlands
[5]Synthetic Biology and Cell Engineering, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, Groningen, The Netherlands

*Present address: Hazrat Ali, Industrial Biotechnology Division, National Institute for Biotechnology and Genetic Engineering, Faisalabad, Pakistan; Yvonne Nygård, Department of Biology and Biological Engineering, Industrial Biotechnology Division, Chalmers University of Technology, Gothenburg, Sweden.

Applied and Environmental Microbiology, 2018, 84(4).

ABSTRACT

Chrysogine is a yellow pigment produced by *Penicillium rubens* and other filamentous fungi. Although it was first isolated in 1973, the biosynthetic pathway has so far not been resolved. Here, we show that the deletion of the highly expressed non-ribosomal peptide synthetase (NRPS) gene *Pc21g12630* (*chyA*) resulted in a loss in the production of chrysogine and thirteen related compounds in the culture broth of *P. rubens*. Each of the genes of the *chyA*-containing gene cluster were individually deleted and corresponding mutants were examined by metabolic profiling in order to elucidate their function. The data suggest that the NRPS ChyA mediates the condensation of anthranilic acid and alanine into the intermediate 2-(2-aminopropanamido)benzoic acid, which was verified by feeding experiments of a *ΔchyA* strain with the chemically synthesized product. The remainder of the pathway is highly branched yielding at least thirteen chrysogine related compounds.

IMPORTANCE

Penicillium rubens is used in industry for the production of β-lactams, but also produces several other secondary metabolites. The yellow pigment chrysogine is one of the most abundant metabolites in the culture broth next to β-lactams. Here, we have characterized the biosynthetic gene cluster involved in chrysogine production and elucidated a complex and highly branched biosynthetic pathway assigning each of the chrysogine cluster genes to biosynthetic steps and metabolic intermediates. The work further unlocks the metabolic potential of filamentous fungi and the complexity of secondary metabolite pathways.

INTRODUCTION

Penicillium rubens and several other filamentous fungi produce the yellow pigment chrysogine (1, 2). Pigments are known to protect the microorganism against adverse environmental conditions, such as UV radiation or harmful organisms (3). The function of chrysogine has not been extensively investigated, but it has been shown that it has no antimicrobial nor anticancer activity (4). *N*-pyruvoylanthranilamide (2-(2-oxopropanamido)benzamide), a related compound produced by *P. rubens* (5) and also identified in *Colletotrichum lagenarium*, has instead anti auxin activity (6).

Chrysogine was first isolated in 1973 by Hikino et al. (5), who observed an increased production upon feeding with anthranilic acid and pyruvic acid. The putative biosynthetic gene cluster has been identified in *P. rubens* (7, 8) and includes a non-ribosomal peptide synthetase (NRPS). Recently Wollenberg et al. showed that a dimodular NRPS is responsible for chrysogine biosynthesis in *Fusarium graminearum* and also suggested a putative cluster (9) homologous to the respective gene cluster of *P. rubens*. However, the actual biosynthetic pathway has remained elusive.

NRPSs are complex multi-modular enzymes that use amino acids and carboxylic acids as substrates (10). The genome of *P. rubens* contains ten genes that encode NRPSs (11). Nonetheless, transcriptomic analysis performed on chemostat cultures of *P. rubens* Wisconsin 54-1255 and the industrially improved DS17690 strain showed that only four of these *NRPS* genes are expressed (11). This set includes three *NRPS* genes that are respectively involved in the biosynthesis of penicillins (12), roquefortines (13) and hydrophobic cyclic tetrapeptides (14). The fourth highly expressed *NRPS* (7–9) is therefore potentially involved in the biosynthesis of chrysogine, that is among the most abundant secondary metabolites produced by this fungus. Furthermore, five genes flanking *Pc21g12630* are also highly co-expressed, suggesting they form a gene cluster (11).

Here, by overexpression and deletion of the core *NRPS* gene of the chrysogine pathway, deletion of the individual pathway gene and by feeding experiments using chemically synthesized intermediates, we elucidate a complex and branched pathway of at least thirteen compounds, assigning a function to each enzyme of the biosynthetic gene cluster.

MATERIALS AND METHODS

Fungal strains, media and culture conditions

P. rubens DS68530 was kindly provided by DSM Sinochem Pharmaceuticals. DS68530 lacks the penicillin gene cluster and the *hdfA* gene (15, 16). For RNA extraction and metabolite analysis, strains were pre-grown in YGG medium (17) for 24 hours. Next, 3 ml of culture inoculum was transferred into 22 ml of secondary metabolites production (SMP) medium (13) and growth was continued for the time indicated. The *Pc21g12630* (*chyA*) overexpression strain was grown in SMP medium, lacking urea and CH_3COONH_4, and supplemented with 2 g/L acetamide for plasmid maintenance. The Δ*chyA* strain was fed with 300 µM of compound **A** or **B** after 48 h of growth. All cultivations were performed as 25 ml cultures in 100 ml erlenmeyer flasks shaken at 200 rpm and 25°C.

Construction of deletion and overexpression plasmids

Plasmids for the deletion of the chrysogine genes were built by PCR amplification of 1–2 kbp of the 5' and 3' flanking regions of each gene, using gDNA from the DS68530 strain as template. All primers used in this study are listed in Tables 1 and 2, the constructed plasmids are shown in the supplementary material.

For the deletion of *Pc21g12630 (chyA)*, *Pc21g12570* (*chyE*), *Pc21g12590* (*chyH*), *Pc21g12610* (*chyM*) and *Pc21g12640* genes, the Multisite Gateway® Three-Fragment Vector Construction Kit (Invitrogen) was used. PCR products were inserted into the donor vectors pDONR4-R1 and pDONR2-R3 by the BP clonase II™ reaction. The resulting plasmids were mixed with the vector carrying the selection marker (pDONR-amdS or pDONR-phleo), the destination vector pDESTR4-R3 and the LB clonase II™ mixture, to form the final constructs. The acetamidase gene *amdS* (17, 18) was employed as a marker for the deletion of *chyH*, *chyM*, *Pc21g12640* genes, while the phleomycin resistance gene was used for selecting *chyA* and *chyE* deleted strains. The modular cloning (MoClo) system (19) was used for building *Pc21g12600* (*chyC*) and *Pc21g12620* (*chyD*) deletion vectors containing an *amdS* marker cassette.

Due to its strength, the *pcbC* promoter was chosen for overexpression of *chyA*, followed by the *penDE* terminator. All genetic elements were amplified from *P. rubens* DS68530 gDNA and the *chyA* expression cassette was built in subsequent steps of digestions and ligation, using

pCM251 (Euroscarf) as backbone vector. The promoter and terminator were digested with *BamHI*, *PmeI* and *NotI* enzymes for cloning into pCM251. *ChyA* was inserted into the resulting pCM251 plasmid after digestion with *AscI* and *PmeI*. The expression cassette was digested with *NotI* for the insertion into pDSM-JAK108 (20), to form pDSM108_AV1. pDSM-JAK108 contains the AMA1 (autonomous maintenance in *Aspergillus*) (21) sequence, the *dsRed* gene for visualization of the cells and the essential gene *tif35*. In this study the *tif35* gene on the plasmid was replaced with an *amdS* cassette by *in vivo* homologous recombination in *P. rubens*. The *amdS* cassette containing 100 bp flanks homologous to pDSM108_AV1 was obtained by oligonucleotide extension-PCR, using pDONR-amdS as template.

Table 1. Oligonucleotide primers used for amplifying the 5' and 3' flanking regions of the targeted genes and for qPCR.

Primer name	Sequence (5'-3')
chyA_5'_fw	GGGGACAACTTTGTATAGAAAAGTTGGGTACCGTTCGTACAC ACCATTCCGGCTG
chyA_5'_rv	GGGGACTGCTTTTTTGTACAAACTTGCATCGATCCTTGATGCC TACAGC
chyA_3'_fw	GGGGACAGCTTTCTTGTACAAAGTGGAAGAGATTGCGAGAGT TGGCTGG
chyA_3'_rv	GGGGACAACTTTGTATAATAAAGTTGGGTACCACTCGAAGGC TCCGTTCTCGGC
chyC_5'_fw	TTGAAGACAATGCCCCTGCAGGTGGGTCGGTATCACAACGAC CG
chyC_5'_rv	TTGAAGACAATTGCGTCCCGTTCGCATGGTTACATAGCT
chyC_3'_fw	GGGGACAACTTTGTATAATAAAGTTGGGTACCACTCGAAGGC TCCGTTCTCGGC
chyC_3'_rv	TTGAAGACAAACTAGTTGAAGAAGTTGGTGTAGTTTGAGAATG
chyD_5'_fw	TTGAAGACAAGGAGCCTGCAGGGATCTCAAAGACTATTATCA AGGAAAGGA
chyD_5'_rv	TTGAAGACAAAGCGGGGTGTCGCATGATTATATCTATAGT
chyD_3'_fw	TTGAAGACAAGGAGTTTGAGATTGAGATGAAAGGATTTGGAA AG
chyD_3'_rv	TTGAAGACAAAGCGCCTGCAGGCGGGCATCTTCACGATCCAA TAG
chyE_5'_fw	GGGGACAACTTTGTATAGAAAAGTTGCGTGCAGCAAAGACGA CATTCG
chyE_5'_rv	GGGGACTGCTTTTTTGTACAAACTTGAGGTATTGGGAATAGA CCGGCC

Primer name	Sequence (5'-3')
chyE_3'_fw	GGGGACAGCTTTCTTGTACAAAGTGGCAGTATATCTGACGAG GAAGTGGG
chyE_3'_rv	GGGGACAACTTTGTATAATAAAGTTGTCTCCTAGTATCCGACT TCTCCG
chyH_5'_fw	GGGGACAACTTTGTATAGAAAAGTTGGCATCGTAATATGCTCG ATTTGG
chyH_5'_rv	GGGGACTGCTTTTTTGTACAAACTTGAGTCTATATAAGCGCTC GGAGGC
chyH_3'_fw	GGGGACAGCTTTCTTGTACAAAGTGGATGAGAGTGAAAGTGT TCAGTGCG
chyH_3'_rv	GGGGACAACTTTGTATAATAAAGTTGGAAGGACCCCTGAGAC AGAACC
chyM_5'_fw	GGGGACAACTTTGTATAGAAAAGTTGAACTTCGAGTCGCAGT ATGCGG
chyM_5'_rv	GGGGACTGCTTTTTTGTACAAACTTGGGTGTAATGGAACCCAT TGCAAGG
chyM_3'_fw	GGGGACAACTTTGTATAGAAAAGTTGAACTTCGAGTCGCAGT ATGCGG
chyM_3'_rv	GGGGACTGCTTTTTTGTACAAACTTGGGTGTAATGGAACCCAT TGCAAGG
Pc21g12640_5'_fw	GGGGACAACTTTGTATAGAAAAGTTGCAAGAGATTGCCGATA ACATTGTGG
Pc21g12640_5'_rv	GGGGACTGCTTTTTTGTACAAACTTGATGACTGGTCCGAGGT ACTGG
Pc21g12640_3'_fw	GGGGACAGCTTTCTTGTACAAAGTGGATCATGCACGATGTGG TCATATGG
Pc21g12640_3'_rv	GGGGACAACTTTGTATAATAAAGTTGGCGGCCGCAGATTTCT CGACGTCCGATC
chyA_qPCR_fw	GCACAGGCCAAAGTAACACGTCC
chyA_qPCR_rv	CCGAGGGTTTGTGGTGGATGCC
chyC_qPCR_fw	GTAGACGCCGGTGAGACTTTGATCG
chyC_qPCR_rv	CAACCTAAGCGTCTAATTTTCATCGC
chyD_qPCR_fw	GGAATTCGCTGGCTAACTGGTCTCG
chyD_qPCR_rv	GGCATGTGGTAGACGAATTGGAGC
chyE_qPCR_fw	GGCAAGGGAAATGAATCCAGGTGGC
chyE_qPCR_rv	GATAGATGCCGCTTGTTCGGACC
chyH_qPCR_fw	GGTTGTGGAGCTCTACGAGGCTG
chyH_qPCR_rv	CTGGCAGGGCTCGTCGGTC
chyM_qPCR_fw	CCTGCATGCAGCTCCATACGAGC
chyM_qPCR_rv	CCAACAATAGGTGGAAACAGCTCAGAC
Pc21g12640_qPCR_fw	TGTCTCTCTGTGGGCTGTTCTCAG
Pc21g12640_qPCR_rv	CAAGAGTTCTTACGATGCGTGGCTG
actin_qPCR_fw	CGACTACCTGATGAAGATCCTCGC
actin_qPCR_rv	GTTGAAGGTGGTGACGTGGATACC

Transformation and purification procedures

The deletion plasmids (1.5 µg) were linearized and transformed into *P. rubens* DS68530 protoplasts using a standard protocol (27). pDSM108_AV1 (1 µg) was linearized by digestion with *MluI* enzyme and co-transformed with the *amdS* cassette (1 µg). The transformants were plated on respective selective media (T-agar) (17) and grown at 25°C for 5 days. For strain purification, the colonies were transferred to minimal selective solid media (S-agar) and sporulation media (R-agar) (17). Rice batches were prepared for inoculation of conidia and long-term storage.

Analysis of the gene deletion strains

The absence of the deleted genes was verified by PCR, with gDNA isolated from the knockout strains after 48 h of growth, using an adapted yeast gDNA extraction protocol (28). Primers binding outside the homologous flanking regions were used for amplification of the targeted fragment, after which the PCR products were further verified by sequencing (Macrogen, UK). To verify the correct integration of the *amdS* cassette into pDSM108-AV1, colony PCR were performed on red colonies (bearing the AMA1 plasmid as seen by the DsRed marker on the plasmid).

RNA extraction, cDNA amplification and qPCR analysis

Total RNA was isolated from the DS68530 and Δ*Pc21g12640* strains after 48 h of growth in SMP medium, by using the Trizol™ (Invitrogen) extraction method with additional DNAse treatment (Turbo DNA-free™ kit, Ambion). For the cDNA synthesis, 500 ng of RNA were used (iScript™ cDNA synthesis kit, Bio-Rad). The γ-actin gene was used for normalization. The expression levels were measured in technical duplicates with a MiniOpticon™ system (Bio-Rad) using the Bio-Rad CFX™ manager software, which determines the threshold cycle (Ct) values automatically by regression. The SensiMix™ SYBR Hi-ROX kit (Bioline) was used as mastermix for qPCR. The reactions were run as follows: 95°C for 10 min, followed by 40 cycles of 95°C for 15 sec, 55°C for 30 sec and 72°C for 30 sec.

Metabolite profiling

All the strains used were grown in triplicates for metabolite analysis. Samples were collected after 48 h from the *chyA* overexpression strain

and after 48 and 96 h from the deletion mutants and the parental strain. Samples were taken before the feeding of Δ*chyA*, immediately after the feeding and then after 48 h. All the samples from the different experiments were centrifuged for 10 min, after which the supernatant was filtered with 0.2 µm polytetrafluorethylene (PTFE) syringe filters and stored at -80°C. The analysis of secondary metabolites was performed with an Accella1250™ HPLC system coupled with the ES-MS Orbitrap Exactive™ (Thermo Fisher Scientific, CA), following the method described by Salo *et al.* (29).

Table 2. Oligonucleotide primers for amplification of P*pcbC*, *chyA* and T*penDE* for cloning into pDSM-JAK108; amplification of *amdS* cassette for *in vivo* homologous recombination into pDSM108_AV1; check the correct integration of *amdS* cassette into pDSM108_AV1; check the absence of the genes in the knockout strains and amplification of the deletion cassettes into the genome. PCR products were sent for sequencing by using primers *phleo_seq* and *amdS_seq*, in order to check the purity of the strains.

Primer name	Sequence (5'-3')
P*pcbC_fw*	CAGTGGATCCACGCGTGTCTGTCAATGACCAATAATTGG
P*pcbC_rv*	CATGGTTTAAACGGCGCGCCGGTGTCTAGAAAAATAATGGTGAAAAC
chyA_cloning_fw	CATGGGCGCGCCATGGCTGCCCCATCCATATCGC
chyA_cloning_rv	CATGGTTTAAACTTACTCGAGATATTCGCAGACTGTCTCTTC
T*penDE_fw*	TCTGCGAATATCTCGAGTAAGTTTAAACCAATGCATCTTTTGTATGTAGCTTC
T*penDE_rv*	TCACTATAGGGCGAATTGGAGCTCCACCGCGGTGGCGGCCGCGGCCGCTGATATCCTGTCTTCAGTCTTAAGAC
amdS_hom_rec_fw	CTTATTAATTTGATGTAGGTAAGCCCGCCACAAATATATATTTTTACAAGATACCGTGGAAAAACTTCGTGCTATCACAAAACAGTATACAAAAAATAAGTGGATCCCCCGGGCTGCAGG
amdS_hom_rec_fw	TCCCCTCGAGCTTGTCTGTGATTGCGTTTTTTCTAACACTTGTTGTTGCATCCGATCCGTCCCTACCAATTATTGGTCATTGACAGACACGCGTACCGCTCGTACCATGGGTTGAGTGGT
amdS_int_fw	ACAGCGGAAGACAAGCTTCTAATAAGTGTCAGATAGCAAT
amdS_int_rv	GTTGGCTCCCAGAGCAGCGGTGTCTTTCGTATTCAGGCAGCTAAAC
chyA_fw	CCATATCGCCGTTATTTGCC
chyA_rv	GACGGCAACATGTAGGAAAC
chyC_fw	ATGGCCCGCATCCTGATCAC
chyC_rv	TTAAGCTGGGAGCTTAATACCGGTGAT
chyD_fw	ATGTGTGGAATAAGTGCATTTCTGTGTC
chyD_rv	TCAGTTTGGCAGGGCACCAG
chyE_fw	ATGGACTCAGTGAGCAATCTAAAG

Primer name	Sequence (5'-3')
chyE_rv	CTATTCTGACAGCCACTGCAAA
chyH_fw	TCGCGATGCCGACTATAAAG
chyH_rv	GCCCATAGAAGCTGAACATC
chyM_fw	ATGGGTTCCATTACACCCTCGC
chyM_rv	TCACCAGAATGCTGCACACCG
Pc21g12640_fw	ATGTCTTCAGCCCCCGGTCT
Pc21g12640_rv	CTAGAATATGTCATCCTCGGATTGGAACC
actin_fw	ATGGAGGGTATGTTATTCCAGTTGTGG
actin_rv	TGCGGTGAACGATGGAAGGACC
phleo in chyA locus_fw	CAACGCCCACGAGCATCTGGT GCCAGAAACTCGACTCGTGGCTC
phleo in chyA locus_rv	
amdS in chyC locus_fw	TCACCAGAATGCTGCACACCG
amdS in chyC locus_rv	GATACCCCTTAGCCCGTCATCCAAA
phleo in chyE locus_fw	CCATGTCGGGTGTAGATCG
phleo in chyE locus_rv	GCCCATAGAAGCTGAACATC
amdS in chyM locus_fw	CTTGTCAAGTCTGCGACCAGCAC
amdS in chyM locus_rv	ACGAAGAGGCACTCGCGTCAC
amdS in Pc21g12640 locus_fw	CAAACAGATGAAGACTGGGG
amdS in Pc21g12640 locus_rv	GGCTCAAACTTGCGCTTAG
phleo_seq	ATGGCCAAGTTGACCAGTGCCGTT
amds_seq	TCCCCTAAGTAAGTACTTTGCTA

RESULTS

Identification of chrysogine related compounds

In order to identify the secondary metabolites produced by the NRPS Pc21g12630, this gene was deleted from *P. rubens* DS68530 by homologous recombination. In this strain, the penicillin cluster is removed (15, 16), facilitating further identification of other secondary metabolites as the metabolite profile is not dominated by β lactams. The strain deleted of the *Pc21g12630* gene did not produce chrysogine and thirteen

other metabolites, from now on referred to as chrysogine related compounds (Table 3). This identified Pc21g12630 as the NRPS responsible for chrysogine biosynthesis and thus this gene was named *chyA*.

Compounds **1**, **2**, **3**, **4**, **8** and **13** were isolated by preparative HPLC and their structures were determined by NMR (Supplemental material). Compound **1** was confirmed to be chrysogine and **3** was identified as *N*-pyruvoylanthranilamide (2-(2-oxopropanamido)benzamide). These compounds were first described in *P. rubens* by Hikino et al. (5). **2** was found to be *N*-acetylalanylanthranilamide (2-(2-acetamidopropanamido) benzamide), previously isolated from a marine *Penicillium* species (22). **4**, **8** and **13** were identified as novel metabolites that are clearly related to chrysogine. The structures of compounds **14** (2-(2-aminopropanamido)benzoic acid) and **15** (the amidated form of compound **14**, 2-(2-aminopropanamido)benzamidine) were further confirmed by the comparison of their HPLC retention time with those of the independently synthesized standards (Supplemental material), while the structures of compounds **5** and **12** were proposed based on their molecular formula. We could not assign a structure to **6**, **7**, **9** and **10** nor we proceeded with the isolation of these compounds due to their low production.

Transcriptomic analysis performed on chemostat cultures of *P. rubens* Wisconsin 54-1255 and the industrial improved DS17690 strain showed that five genes flanking *chyA* (*Pc21g12570*, *Pc21g12590*, *Pc21g12600*, *Pc21g12610*, *Pc21g12620*) were also highly expressed, indicating that they could be part of the chrysogine gene cluster (11) (Figure 1). Furthermore, quantitative PCR confirmed the expression of the above listed genes in the DS68530 strain after 48 h of growth in a SMP medium (Figure S2). Therefore, we tentatively assigned these as *chy* genes. *Pc21g12640*, found adjacent to the *chy* genes, exhibits a strong similarity with a cutinase transcription factor beta from *Fusarium solani* (11). Although not significantly expressed in DS68530, its possible role as regulator of the cluster was also investigated.

Expression of the *NRPS chyA* in a chrysogine cluster deleted strain

In order to identify the products of the *NRPS chyA*, a chrysogine cluster deleted strain (8) was used to overexpress the *chyA* gene from an episomal AMA1 based plasmid. The *chyA* overexpressing strain produced compounds **14**, **8** and **13** (Figure 2). It is likely that compound **14** is the

Table 3 Production of chrysogine and related metabolites from strain DS68530. Numbers in the DS68530 columns represent the peak areas of the compounds corrected for the internal standard reserpine. The culture broth of strain DS68530 was analyzed after 48 and 96 h of growth in a SMP medium.

compound number	compound name	formula	acquired [M+H]+	RT (min)	DS68530 48 h	DS68530 96 h
1	chrysogine	$C_{10}H_{10}O_2N_2$	191.08	12.02	185,02	236,37
2	2-(2-acetamidopropanamido)benzamide	$C_{12}H_{15}O_3N_3$	250.12	10.77	19,85	53,92
3	2-(2-oxopropanamido)benzamide	$C_{10}H_{10}O_3N_2$	207.08	10.32	0,65	11,26
4	3-((1-((2-carbamoylphenyl)amino)-1-oxopropan-2-yl)amino)-3-oxopropanoic acid	$C_{13}H_{15}O_5N_3$	294.11	10.62	24,84	152,64
5	(1-((2-carbamoylphenyl)amino)-1-oxopropan-2-yl)glutamine	$C_{15}H_{20}O_5N_4$	337.15	8.60	4,92	0
6	chrysogine related	$C_{15}H_{18}O_6N_3$	336.12	9.02	0	2,44
7	chrysogine related	$C_{13}H_{12}O_5N_2$	277.08	11.06	0	0,82
8	2-(2-(2-carboxyacetamido)propanamido)benzoic acid	$C_{13}H_{14}O_6N_2$	295.09	15.20	0	1,12
9	chrysogine related	$C_{20}H_{20}O_6N_4$	413.14	14.95	0,42	0,68
10	chrysogine related	$C_{20}H_{20}O_6N_4$	413.14	15.70	0,42	0,57
12	3-oxo-3-((1-(4-oxo-1,4-dihydroquinazolin-2-yl)ethyl)amino)propanoic acid	$C_{13}H_{13}O_4N_3$	276.10	12.59	0,50	1,10
13	2-(2-((4-amino-1-carboxy-4-oxobutyl)amino)propanamido)benzoic acid	$C_{15}H_{19}O_6N_3$	338.13	12.65	2,91	0,20
14	2-(2-aminopropanamido)benzoic acid	$C_{10}H_{12}O_3N_2$	209.09	6.20	1,15	0,18
15	2-(2-aminopropanamido)benzamide	$C_{10}H_{13}O_2N_3$	208.11	2.47	0,17	0

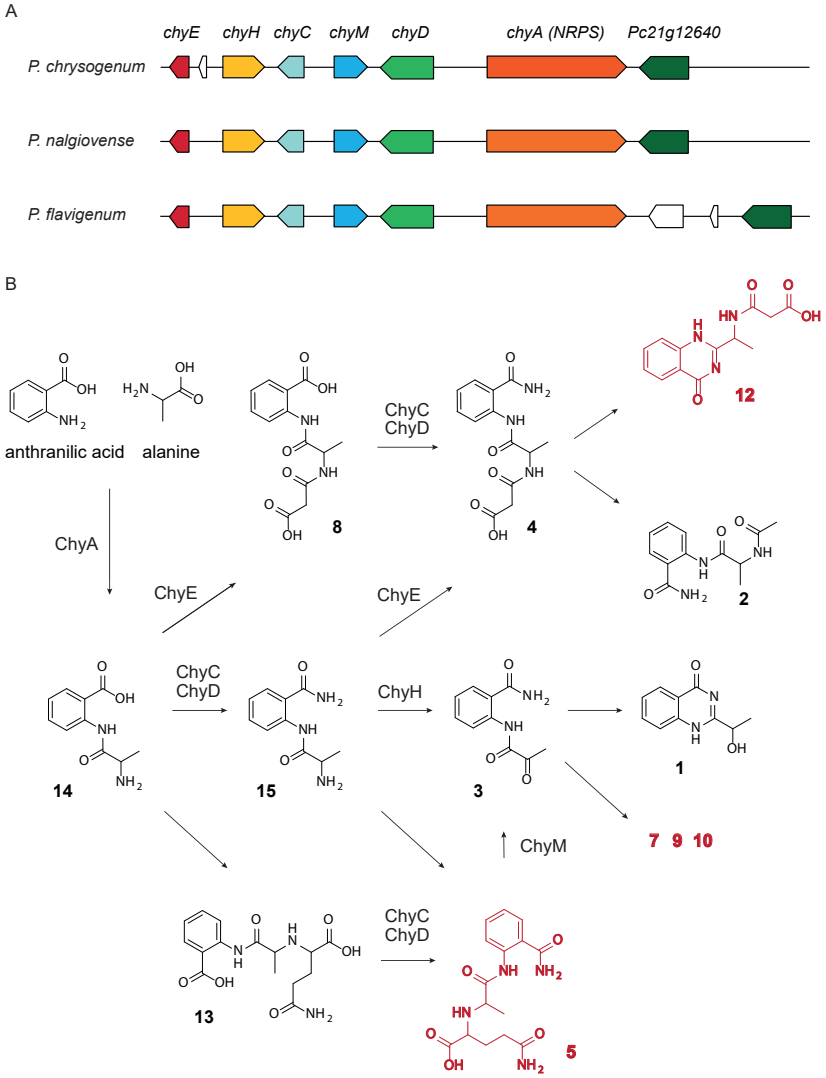


Fig.1 Representation of the chrysogine biosynthetic gene cluster and proposed pathway. The chrysogine biosynthetic gene cluster in *P. rubens* and two other chrysogine-producing species. Genes with the same color have >80% identity. This study identified ChyA as the NRPS, ChyE as malonyl transferase, and ChyD as amidase; ChyC participates in amidation reactions, while ChyH and ChyM are involved in oxidation reactions. The substrates of ChyA and the compounds identified in this study are depicted in black, and the putative structures and uncharacterized compounds are represented in red

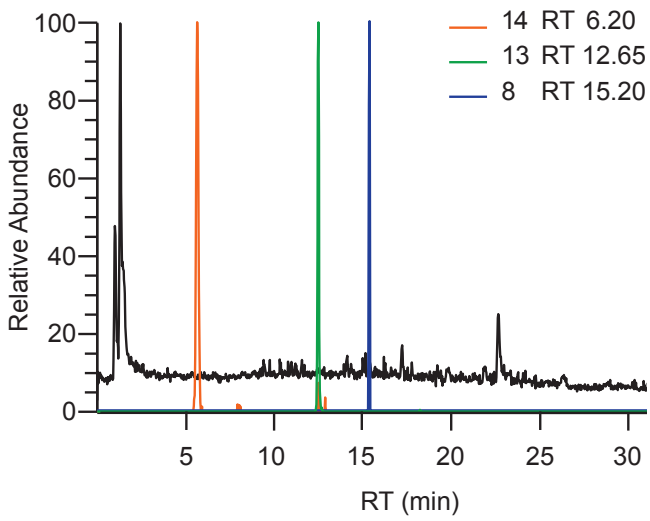


Fig.2 Chromatogram of culture broth from the *chyA*-expressing strain. Total ion chromatogram (TIC; black) and extracted ion chromatograms (EIC; colored) of secondary metabolites produced by the *chyA*-expressing strain after 48 h of growth in an SMP medium.

immediate product of the NRPS and that this compound is derived from the condensation of anthranilic acid and alanine. **8** and **13** could respectively be derived from compound **14** by addition of a malonyl and glutaminyl group. Our data suggest an immediate branching of the pathway, where two groups of compounds are derived from **8** and **13**.

Metabolite profiles of *chy* gene deletion strains

The expression of *chyA* in a chrysogine cluster deleted strain allowed the identification of the product of the NRPS and metabolites produced early in the pathway. To elucidate how the initial products were further modified by the enzymes of the cluster and resolve the complete pathway, individual *chy* genes knockout strains were made and metabolite profiling was performed (Table 4).

The deletion of *chyD* led to a depletion of most chrysogine related metabolites – only compounds **14**, **8** and **13** were accumulated during cultivation of this mutant. This suggests that ChyD is an early enzyme of the pathway, being responsible for converting **14**, **8** and **13** into downstream compounds. Based on its formula, we propose that **14** is converted into **15**, which is its amidated form.

Table 4 Secondary metabolites of the chrysogine pathway in the knockout strains compared to that of the parental strain. Numbers represent the peak areas of the compounds corrected for the internal standard reserpine and relative to those of the parental strain, DS68530. The culture broth of the strains was analyzed after 48 and 96 h of growth in a SMP medium.

	ΔchyC		*ΔchyD*		*ΔchyE*		*ΔchyH*		*ΔchyM*		*ΔPc21g12640*	
compound name	**48 h**	**96 h**	**48 h**	**96 h**	**48 h**	**96 h**	**48 h**	**96 h**	**48 h**	**96 h**	**48 h**	**96 h**
1	0,19	0,37	0	0	0,66	0,73	0,03	0,03	0,01	0,02	0,96	0,83
2	0,16	0,23	0	0	0	0,01	0,86	0,77	1,88	1,89	0,80	0,70
3	0,23	0,21	0	0	0,68	0,58	0	0,02	0	0	0,93	0,82
4	0,19	0,45	0	0	0	0,01	0,84	0,75	1,16	1,31	0,84	0,75
5	0	0	0	0	0,67	0,41	0,66	0	1,44	2,40	0,75	0,04
6	0	0,05	0	0	0	1,03	0	0,51	0	0,42	0	0,58
7	0	0,45	0	0,09	0	0,37	0	0	0	0	0	0,69
8	0	3,86	0,31	48,40	0	0,23	0	1,80	0	1,09	0	0,42
9	0	0	0	0	0,51	0,54	0	0	0	0	0,86	0,50
10	0	0	0	0	0,51	0,92	0	0	0	0	0,86	0,80
12	0,17	0,24	0	0	0	0	0,97	0,48	3,39	3,07	0,93	0,78
13	0	4,09	15,59	12,75	0,52	9,16	0,67	0,22	1,36	40,26	0,81	2,86
14	0,33	15,87	11,56	16,73	0,34	2,34	0,92	0,28	1,98	13,99	0,89	1,24
15	0	0	0	0	0,05	0,01	0,99	0	2,64	0,05	0,69	0

The Δ*chyC* strain showed a metabolite profile similar to the Δ*chyD* strain suggesting that ChyC could be also involved in the conversion of **14**, **8** and **13**. Nonetheless, downstream compounds were still produced in low amount in the Δ*chyC* strain.

In the Δ*chyE* strain, **2**, **4**, **8** and **12** were not detected or produced in low concentrations compared to the parental strain, suggesting that these compounds belong to the same initial branch of the pathway. Based on the structures and molecular formula available, **2**, **4** and **12** are derived from **8**, with **4** being most likely spontaneously converted into **2** and **12**. Since ChyE affected the production of **8** and downstream compounds and accumulated **14** after 96 h of growth, we propose that this enzyme converts **14** into **8**.

A trend opposite to the metabolite profile of Δ*chyE* can be observed in the Δ*chyM* strain. Peak areas of **2**, **4**, **8** and **12** were comparable to DS68530 strain, while **1**, **3**, **7**, **9** and **10** were absent or detected in low amounts. This indicates that these compounds are part of an independent branch of the pathway and derived from **13**. The result is confirmed by the accumulation of **14** and **13** in the Δ*chyM* strain. The molecular formula of **5** suggests it is derived from **13** and that it is the precursor of **3**, which is further converted into **1**, **7**, **9** and **10**. Because **3** and downstream compounds were not produced in this mutant, we propose that ChyM is responsible for the conversion of **5** into **3**. Chrysogine (**1**) is likely formed by a spontaneous ring closure from **3**. Compounds **9** and **10** are isomers, having the same molecular mass but different retention times on HPLC.

Finally, the Δ*chyH* strain showed a metabolite profile similar to that of Δ*chyM*, suggesting that both the enzymes are needed for the formation of the same compounds. Nonetheless, Δ*chyH* did not accumulate **13** and **5**, suggesting that ChyH forms **1**, **3**, **7**, **9** and **10** through an independent path. In the analysis of the mutant strains, we could not assign the position of compound **6** in the pathway. Based on the molecular formula, **6** could be an unstable precursor of **13**.

Metabolite profile and gene expression in a strain with a deletion of a putative transcription factor

Pc21g12640 encodes a putative transcription factor and, because of its chromosomal location in the vicinity of the chrysogine biosynthetic gene cluster, it would be plausible that it acts as a local regulator of

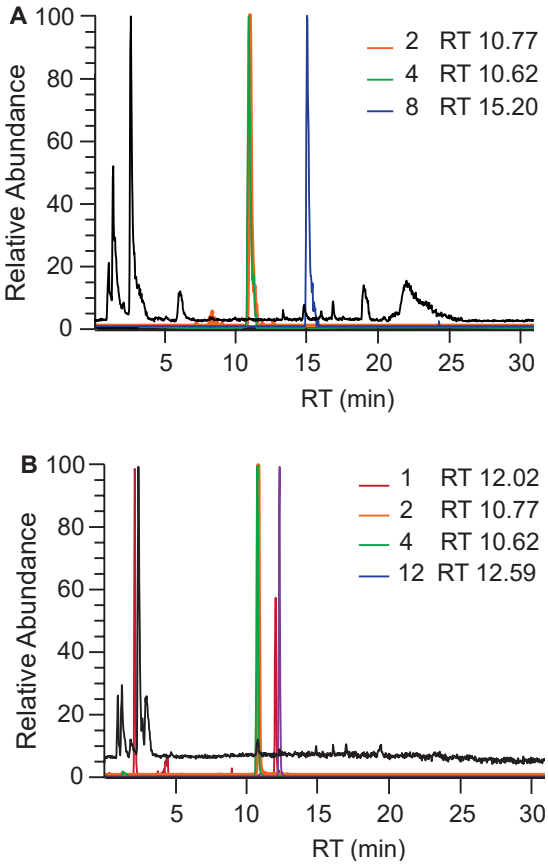


Fig.3 Chromatogram of culture broth from the Δ*chyA* strain fed with compound 14 or 15. TIC (black) and EIC (colored) of secondary metabolites produced by the Δ*chyA* strain fed with compound 14 (A) or 15 (B) after 48 h from the feeding.

this pathway. Although *Pc21g12640* is not significantly expressed in the DS68530 strain, transcription factors can regulate transcription even when present at very low levels. Therefore, to investigate its possible role as a regulator of the chrysogine cluster, *Pc21g12640* was deleted from strain DS68530. Nonetheless, the Δ*Pc21g12640* strain did not show any significant changes in the chrysogine related metabolite profiles compared to the parental strain (Table 4). Similarly, qPCR indicated that the deletion of *Pc21g12640* did not significantly affect the expression of the genes of the chrysogine cluster (Figure S2). Thus, *Pc21g12640* is not part of the chrysogine biosynthetic gene cluster.

Feeding of the Δ*chyA* strain with compounds 14 and 15

In order to further investigate the role of compounds **14** and **15** as potential NRPS products, the Δ*chyA* strain was fed with chemically synthesized variants of these. Based on the formula, **15** is the amidated form of **14**.

Above we showed that the expression of *chyA* in the chrysogine deleted strain resulted in the production of **14**, **8** and **13**. The Δ*chyA* strain fed with **14** produced **2**, **4** and **8**, while **13** and downstream compounds were not detected (Figure 3A). This result suggests that the conversion of **14** into **8** is faster than its conversion into **13**. The feeding with **15** resulted in the production of metabolites that are derived from **8** (**2**, **4**, **12**) and **13** (**1**, Figure 3B). As compound **15** is very similar to compound **14**, we suggest that **15** undergoes the same reactions, being converted into **4** by ChyE and into **5** by a transaminase. Since Δ*chyH* affected the production of **3** and downstream metabolites without any accumulation of **5**, we propose that ChyH is involved in the biosynthesis of **3** from **15**. Therefore, the late metabolites can be formed from two different paths.

Distribution and diversity of chrysogine gene clusters in *Penicillia* species

Since the above studies characterized the chrysogine biosynthetic gene cluster, the distribution of this gene cluster in other *Penicillia* species was investigated (Figure 1). The *chy* genes and *Pc21g12640* from *P. rubens* were blasted against the genomes of two known chrysogine producers (2), *P. nalgiovense* and *P. flavigenum*, recently sequenced by Nielsen et al. (23). These genomes contain a chrysogine gene cluster with similar gene organization, while a *Pc21g12580* homolog is missing, supporting the notion that this gene is not essential for chrysogine biosynthesis. Interestingly, *P. flavigenum* has two extra genes nearby the *NRPS* gene, suggesting that it may produce additional chrysogine related metabolites.

DISCUSSION

Chrysogine was isolated from the culture broth of *P. rubens* in 1973 (5) and found to be produced also by other filamentous fungi (1, 2). Chrysogine biosynthesis is mediated by a dimodular NRPS that we

recently identified in *P. rubens* (7, 8) and that was also shown to be responsible for chrysogine biosynthesis in *Fusarium graminearum* (9). Although the biosynthetic gene cluster was suggested, the role of the enzymes in the pathway has sofar not been characterized. In this work, we assigned a function to each enzyme of the cluster and elucidated a complex pathway, validating the compound structures by NMR. The pathway is highly branched, with some enzymes involved in multiple steps of the biosynthesis (Figure 1).

The NRPS ChyA is a 260 kDa dimodular enzyme which is predicted to contain two adenylation domains. The increased production of chrysogine upon feeding with anthranilic acid and pyruvic acid (5) suggests these molecules are possible substrates of the NRPS. However, here we identify compound **14** as the direct product of ChyA, showing that the NRPS in addition to anthranilic acid utilizes alanine instead of pyruvic acid. However, alanine is readily derived from pyruvic acid by transamination which explains why pyruvic acid stimulates chrysogine production.

Compound **14** acts as a substrate for several enzymes, which immediately results in a split in the pathway by forming **8**, **13** and **15**, the latter being the amidated form of compound **14**. Two independent groups of compounds are derived from **8** and **13**. Since **15** undergoes the same reactions as **14**, the more distal metabolites in the pathway can be formed via either branch that converge.

Transcriptomic data (11) suggested that *chyA* and five flanking genes could form a cluster. These genes are co-expressed under a set of conditions, whereas expression profiles in the flanking regions of the putative gene cluster vary. Metabolic profiling of the mutant strain indicated that ChyE is a malonyl transferase, which can convert **14** and **15** into **8** and **4**, respectively. Interestingly, the expression of *chyA* in a chrysogine cluster deleted strain showed that **14** can be converted into **8** without involvement of any of the enzymes of the cluster; this conversion likely involves a transferase. In line with this observation, the deletion of *chyE* did not lead to a complete depletion of **8** and downstream metabolites, although it significantly decreased the amounts produced. These data suggest that *chyE* is part of the biosynthetic cluster, as it is co-expressed together with the other genes (11) and its deletion affects chrysogine metabolites production, but one or more other transferases can catalyze the same reactions. The orthologous gene in *Fusarium* species is not

involved in chrysogine biosynthesis, showing a different expression pattern compared to the genes of the cluster (9).

Also compound **13** was formed by the strain that solely expresses *chyA*, likely through the involvement of a transaminase, which is not part of the gene cluster. Based on sequence alignment, no genes encoding for a transaminase have been identified in the immediate vicinity of the chrysogine genes, but the genome contains many transaminases.

Our data indicate that ChyD is an amidase, being responsible for the amidation of the carboxylic acid moiety of **14**, **8** and **13**, in line with the bioinformatics prediction of ChyD as an asparagine synthetase, which amidates aspartate to form asparagine. The Δ*chyC* strain showed a metabolite profile similar to that of the Δ*chyD* strain, suggesting that ChyC is involved in the same reactions as ChyD. Indeed, downstream compounds were still produced in low amount in the Δ*chyC* strain. For this reason, we speculate that ChyC plays a more minor role in the amidation reactions compared to ChyD, whose deletion abolished completely the production of the late metabolites. Protein alignment does not provide sufficient information to assign a specific function to ChyC. ChyH and ChyM are predicted to be involved in oxidation reactions and form compound **3** from **15** and **5**, respectively. **3** originates from two further branches in the pathways, yielding chrysogine and **7**, **9** and **10**.

Regulatory genes are usually clustered with secondary metabolite biosynthetic genes (24). Therefore, we hypothesized that the putative transcription factor Pc21g12640 can regulate the expression of the chrysogine genes, since *Pc21g12640* is located downstream of *chyA*. Nonetheless, metabolite profiling and qPCR of the deletion strain gave no indications that Pc21g12640 is involved in the regulation of the *chy* genes. This conclusion is supported by the absence of the transcription factor in *Fusarium* and the other filamentous fungi investigated by Wollenberg et al. (9), although the orthologous gene is present in the genome of other *Penicillia* species (Figure 1).

As already shown for some other fungal secondary metabolites clusters (24, 25), it is possible that the chrysogine biosynthetic genes are regulated by other transcription factors. Moreover, epigenetic regulation has been suggested for the chrysogine cluster. Shwab et al. (26) first demonstrated that secondary metabolites genes can be regulated by chromatin remodeling, for example by histone acetylation. In *P. rubens* DS68530, the deletion of the histone deacetylase *hdaA* resulted in a

significant downregulation of the *chy* genes expression and subsequent reduction of chrysogine biosynthesis (Guzman, Salo and Samol, unpublished data).

Secondary metabolite pathways can provide a wide range of compounds from the initial scaffold molecule. Moreover, the same compounds can be produced through different paths. Branched secondary metabolite pathways have been described before in *P. rubens* (13). The chrysogine pathway is even more branched than the previously described roquefortine pathway, and in this case, chrysogine is the final product of one ramification. As a pigment, chrysogine could contribute to protect the cell from UV light. No antimicrobial activity has been found for this metabolite (4) nor for *N*-acetylalanylanthranilamide (**2**), which was also identified in a marine fungus (22). The function of the other metabolites in the cell remains unknown. Nonetheless, the approaches used in this work and the established methods can provide a blueprint for the elucidation of novel secondary metabolite pathways that potentially specify unknown bioactive compounds. Moreover, the understanding of the biosynthetic mechanisms can help to develop new molecules by feeding with chemically modified intermediates.

ACKNOWLEDGMENTS

The research was supported by a grant from the People Programme (Marie Curie Actions) of the European Union's Seventh Framework Programme FP7/2007-2013, under grant 607332. The financial support from the Dutch Organization for Scientific Research (NWO VIDI grant 723.014.001 for W.S.) is gratefully acknowledged. We thank DSM Sinochem Pharmaceuticals (Delft, the Netherlands) for kindly providing the DS68530 strain.

SUPPLEMENTARY DATA

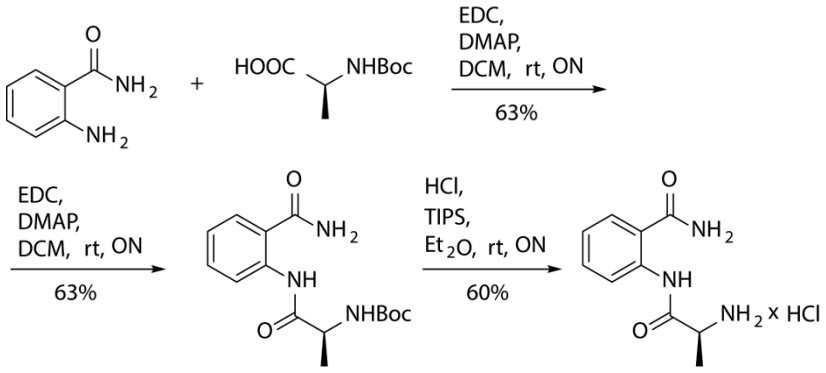
NMR analysis

Samples of compound **3**, **8** and **13** were dissolved in 0.300 ml DMSO, next 0.300 ml of $CDCl_3$ were added and the solution was transferred to a 5 mm NMR tube. NMR spectra were recorded on an Agilent

Technologies 400-MR (400/54 Premium Shielded) spectrometer (400 MHz), Bruker Ascend 700 MHz NMR spectrometer or on a Bruker Ascend 600 MHz NMR spectrometer at 300 K and at low temperature (260 K) with water suppression by means of the standard Bruker pulse program zgcppr. An inter pulse delay of 10 s was chosen for the ^{1}H spectra to ensure quantitative comparison of signal integrals. All ^{13}C-NMR spectra are 1H-broadband decoupled.

COSY, TOCSY, HSQC and HMBC spectra for assignments of signals were recorded with standard Bruker pulse sequences. Chemical shifts are expressed relative to:
In DMSO-d_6: ^{1}H δDMSO = 2.55, ^{13}C δDMSO = 39.5;
In $CDCl_3$: ^{1}H δTMS = 0.00, ^{13}C δ$CDCl_3$ = 77.0;
In MeOD: ^{1}H δMeOH = 3.31, ^{13}C δMeOD = 49.0;

Synthesis of compound 14 and 15

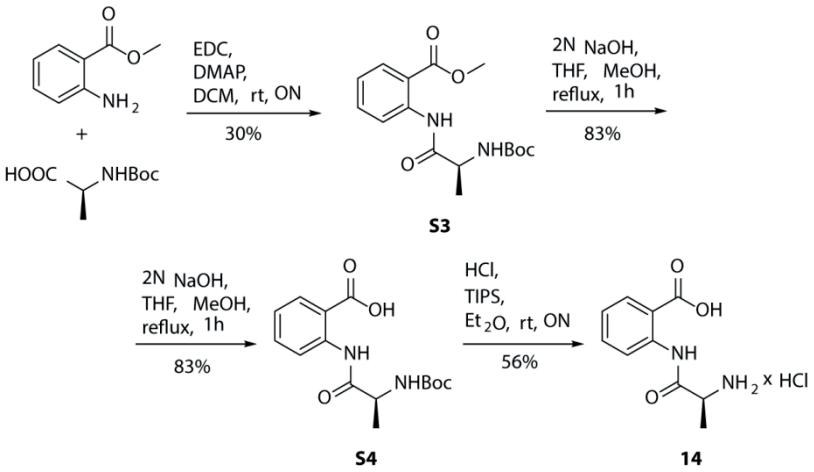


S1 **15**

S1: *tert*-Butyl (*S*)-(1-((2-carbamoylphenyl)amino)-1-oxopropan-2-yl)carbamate. A solution of anthranilamide (1.00 mmol, 126 mg), *N*-Boc-L-alanine (1.00 mmol, 189 mg) and dimethylaminopyridine (0.2 mmol, 25 mg) in DCM (4 mL) was stirred at RT. EDC (1.10 mmol, 210 mg) was added in one portion and the stirring continued overnight. The reaction mixture was diluted with ethyl acetate (80 mL) and washed with 1N HCl (3 × 60 mL), sat. aq. $NaHCO_3$ (3 × 60 mL) and brine (60 mL). The organic phase was dried ($MgSO_4$) and the solvent volume was reduced. Addition of pentane resulted in precipitation of product as a white powder (192 mg, 63 %). R*f* = 0.56 (pentane / AcOEt, 1:1, v/v); Mp. 76-78°C; ^{1}H NMR (400 MHz, DMSO-d_6): δ 1.26 (d, ^{3}J = 7.2 Hz, 3H, CH_3CH), 1.39 (s,

9H, (CH_3)C), 3.85-3.96 (m, 1H, CH_3CH), 7.09 (app t, ^3J = 7.6 Hz, 1H, ArH), 7.45 (br s, 1H, NH), 7.46 (app t, ^3J = 7.6 Hz, 1H, ArH), 7.56 (br s, 1H, NH), 7.76 (d, ^3J = 8.0 Hz, 1H, ArH), 8.18 (br s, 1H, NH), 8.51 (d, ^3J = 8.4 Hz, 1H, ArH), 11.99 (s, 1H, NH); ^{13}C NMR (100 MHz, DMSO-d_6): δ 17.8, 28.7, 52.1, 78.8, 120.3 ,120.3 122.8, 129.0, 132.5, 139.8, 155.8, 170.8, 172.6; HRMS (ESI+) calc. for $[M+H]^+$ ($C_{15}H_{22}N_3O_4$): 308.1604, found: 308.1606.

15: (*S*)-2-(2-aminopropanamido)benzamide hydrochloride. To a solution of compound **S1** (0.55 mmol, 170 mg) and tri-*iso*-propylsilane (0.60 mmol, 123 µL) in Et_2O (10 mL) was added 2M HCl in Et_2O (10 mL). The resulting solution was stirred at RT overnight. A precipitate was formed, which was filtered and washed with excess Et_2O to give, after drying *in vacuo*, a white powder (80 mg, 60%). ^{1}H NMR (400 MHz, CD_3OD): δ 1.63 (d, ^3J = 7.2 Hz, 3H, CH_3CH), 4.16 (q, ^3J = 7.2 Hz, 1H, CH_3CH), 7.21 (app t, ^3J = 7.6 Hz, 1H, ArH), 7.52 (app t, ^3J = 7.6 Hz, 1H, ArH), 8.00 (d, ^3J = 8.0 Hz, 1H, ArH), 8.38 (d, ^3J = 8.0 Hz, 1H, ArH); ^{13}C NMR (100 MHz, CD_3OD): δ 15.7, 49.9, 120.8, 121.2, 123.7, 128.8, 132.2, 138.2, 167.4, 171.9; HRMS (ESI+) calc. for $[M+H]^+$ ($C_{10}H_{14}N_3O_2$): 208.1080, found: 208.1081.



S3: Methyl (*S*)-2-(2-((tert-butoxycarbonyl)amino)propanamido)benzoate. A solution of methyl 2-aminobenzoate (2.00 mmol, 302 mg) and *N*-Boc-L-alanine (2.00 mmol, 378 mg) in DCM 8 mL) was stirred at rt. EDC (2.20 mmol, 421 mg) was added in one portion and the stirring continued overnight. The reaction mixture was diluted with ethyl acetate (80 mL) and washed with 1N HCl (3 × 60 mL), sat. aq. $NaHCO_3$

(3 × 60 mL) and brine (60 mL). The organic phase was dried ($MgSO_4$) and the solvent volume was reduced. Addition of pentane resulted in precipitation of product as a white powder (190 mg, 30 %). Mp. 114-115°C; ^{1}H NMR (400 MHz, $CDCl_3$): δ 1.45 (s, 9H, $(CH_3)C$), 1.47 (d, ^3J = 7.2 Hz, 3H, CH_3CH), 3.89 (s, 3H, CH_3O), 4.25-4.40 (m, 1H, CH_3CH), 5.14 (br s, 1H, NHBoc), 7.07 (app t, ^3J = 8.0 Hz, 1H, ArH), 7.52 (app t, ^3J = 8.0 Hz, 1H, ArH), 8.00 (d, ^3J = 8.0 Hz, 1H, ArH), 8.69 (d, ^3J = 8.4 Hz, 1H, ArH), 11.50 (s, 1H, NH); ^{13}C NMR (100 MHz, $CDCl_3$): δ 18.9, 28.3, 51.6, 52.3, 80.0, 115.3, 120.3, 122.7, 130.8, 134.6, 141.1, 155.2, 168.4, 171.8; HRMS (ESI+) calc. for $[M+H]^+$ ($C_{16}H_{23}N_2O_5$): 323.1602, found: 323.1599.

S4: (*S*)-2-(2-((tert-butoxycarbonyl)amino)propanamido)benzoic acid. To a solution of compound **S3** (0.50 mmol, 156 mg) in MeOH (12 mL) and THF (12 mL) was added 2N aq. NaOH (4 mL). The resulting mixture was heated at reflux for 1 h. The volatiles were evaporated, the residue was redissolved in with ethyl acetate (30 mL) and washed with 1N HCl (2 × 20 mL) and brine (20 mL). The organic phase was dried ($MgSO_4$) and the solvent volume was reduced. Addition of pentane resulted in precipitation of product as a white powder (128 mg, 83 %). Mp. 154-156°C; ^{1}H NMR (400 MHz, DMSO-d_6): δ 1.28 (d, ^3J = 7.2 Hz, 3H, CH_3CH), 1.38 (s, 9H, $(CH_3)C$), 3.90-4.02 (m, 1H, CH_3CH), 7.12 (app t, ^3J = 7.6 Hz, 1H, ArH), 7.51 (d, ^3J = 6.4 Hz, 1H, NHBoc), 7.57 (app t, ^3J = 7.6 Hz, 1H, ArH), 7.97 (d, ^3J = 8.0 Hz, 1H, ArH), 8.61 (d, ^3J = 8.0 Hz, 1H, ArH), 11.65 (s, 1H, NH); ^{13}C NMR (100 MHz, DMSO-d_6): δ 17.8, 28.6, 52.2, 78.9, 116.5, 119.8, 123.0, 131.6, 134.6, 141.2, 155.9, 169.6, 172.9; HRMS (ESI+) calc. for $[M+H]^+$ ($C_{15}H_{21}N_2O_5$): 309.1445, found: 309.1440.

14: (*S*)-2-(2-aminopropanamido)benzoic acid hydrochloride. To a solution of compound **S4** (0.39 mmol, 120 mg) and tri-*iso*-propylsilane (0.50 mmol, 108 µL) in Et_2O (5 mL) was added 2M HCl in Et_2O (5 mL). The resulting solution was stirred at rt overnight. A precipitate was formed, which was filtered and washed with excess Et_2O to give, after drying *in vacuo*, a white powder (53 mg, 56%). ^{1}H NMR (400 MHz, CD_3OD): δ 1.66 (d, ^3J = 7.2 Hz, 3H, CH_3CH), 4.23 (q, ^3J = 7.2 Hz, 1H, CH_3CH), 7.21 (app t, ^3J = 7.6 Hz, 1H, ArH), 7.59 (app t, ^3J = 7.6 Hz, 1H, ArH), 8.10 (d, ^3J = 8.0 Hz, 1H, ArH), 8.51 (d, ^3J = 8.4 Hz, 1H, ArH); ^{13}C NMR (100 MHz, CD_3OD): δ 15.7, 50.0, 116.7, 120.3, 123.4, 131.2, 133.9, 140.0 167.6, 169.9; HRMS (ESI+) calc. for $[M+H]^+$ ($C_{10}H_{13}N_2O_3$): 209.0921, found: 209.0916.

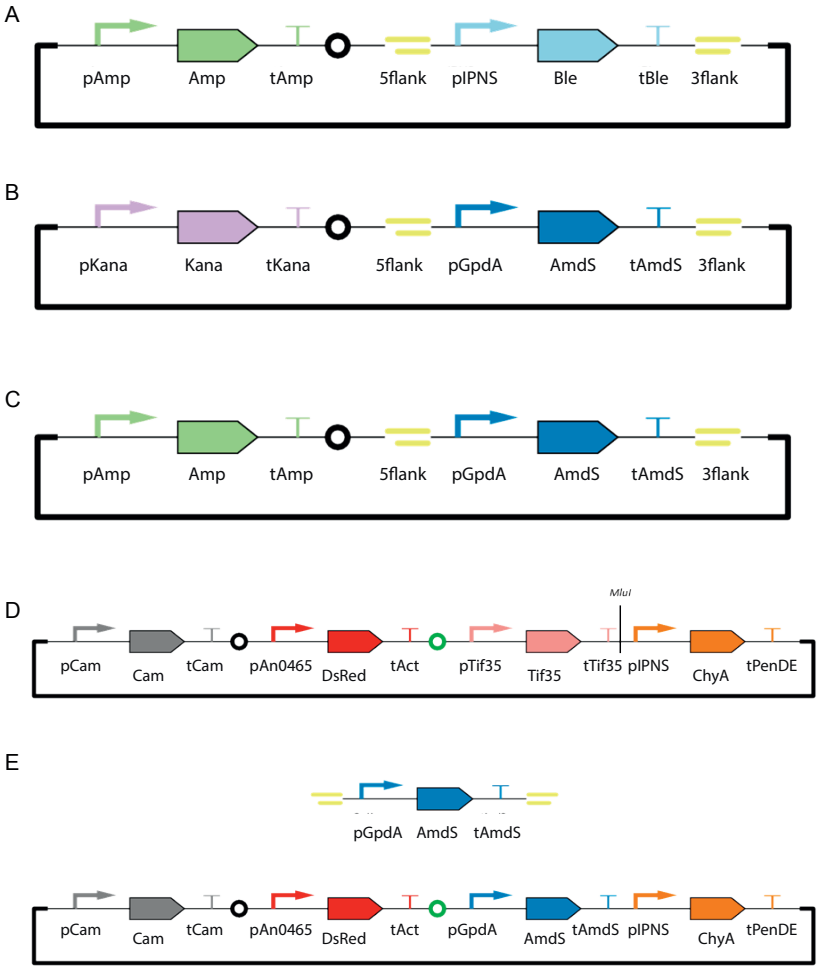


Figure S1. SBOL (Synthetic Biology Open Language) presentation of deletion plasmids for *chyA*, *chyE* (A), *chyC*, *chyD* (B), *chyH*, *chyM* and *Pc21g12640* (C). SBOL presentation of pDSM108_AV1 overexpressing *chyA* and *in vivo* repair cassette (D). *In vivo* recombined plasmid from D (E).

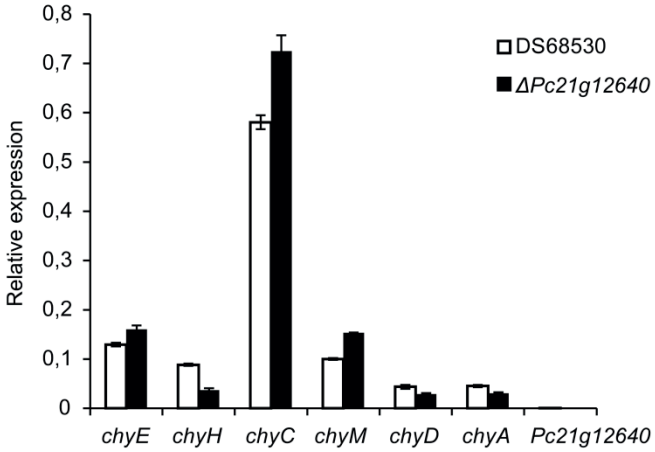


Figure S2. Expression of the putative chrysogine gene cluster in DS68530 and Δ*Pc21g12640* strains. RNA was isolated after 48 h of growth in a SMP medium. Data are expressed relative to actin and represented as mean ± SEM.

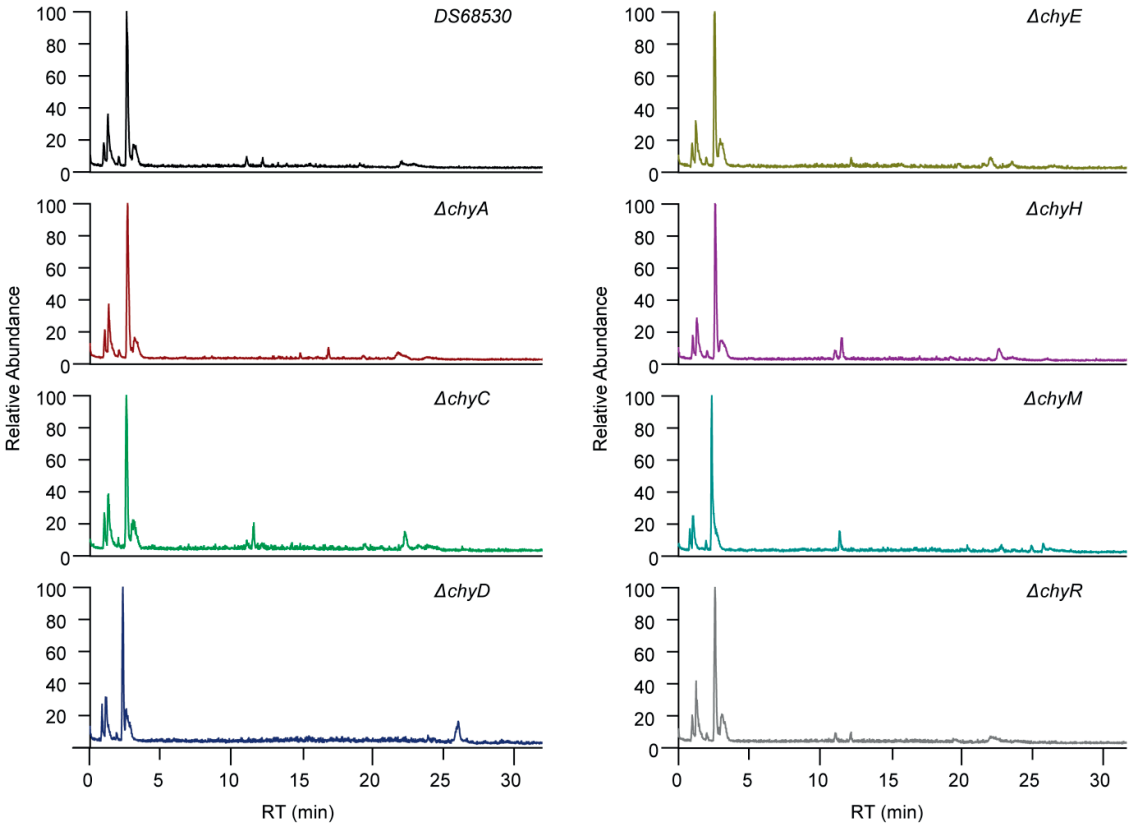


Figure S3. Chromatograms of culture broth from DS68530 and knockout strains after 96 h of growth in a SMP medium.

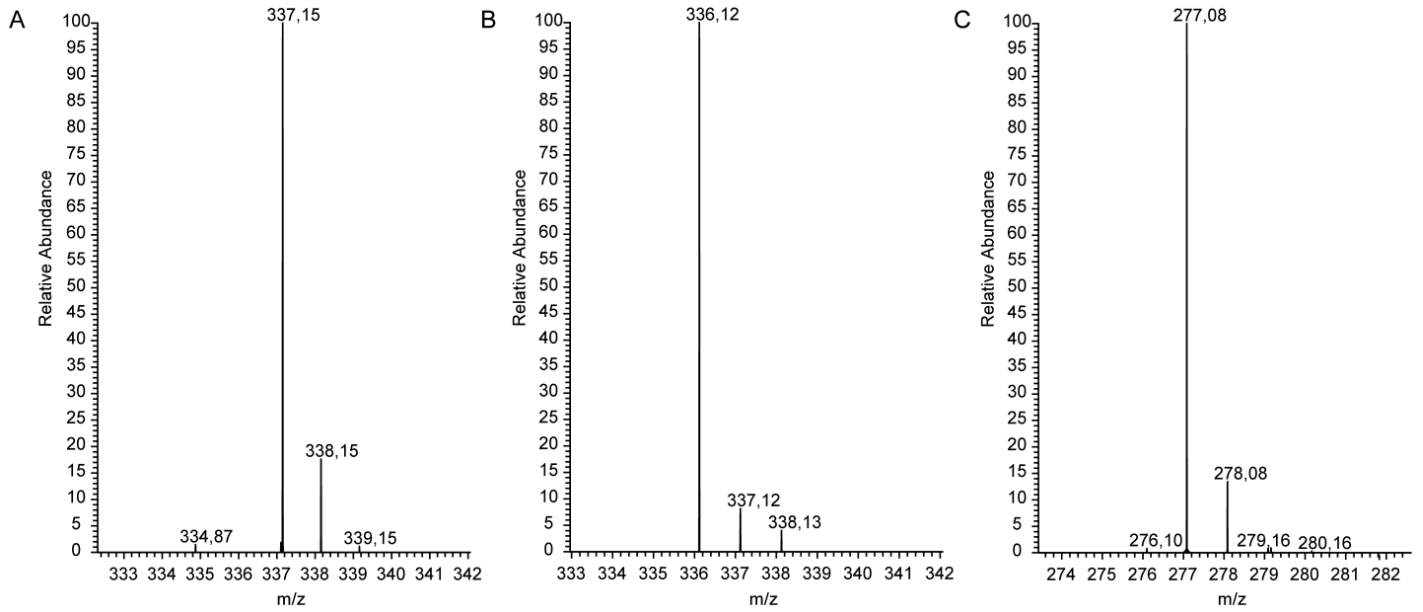


Figure S4. Mass spectra of the uncharacterized chrysogine related compounds. Compound **5** (A), compound **6** (B), compound **7** (C), compound **9** and **10** (D), compound **12** (E) – continues on page 76.

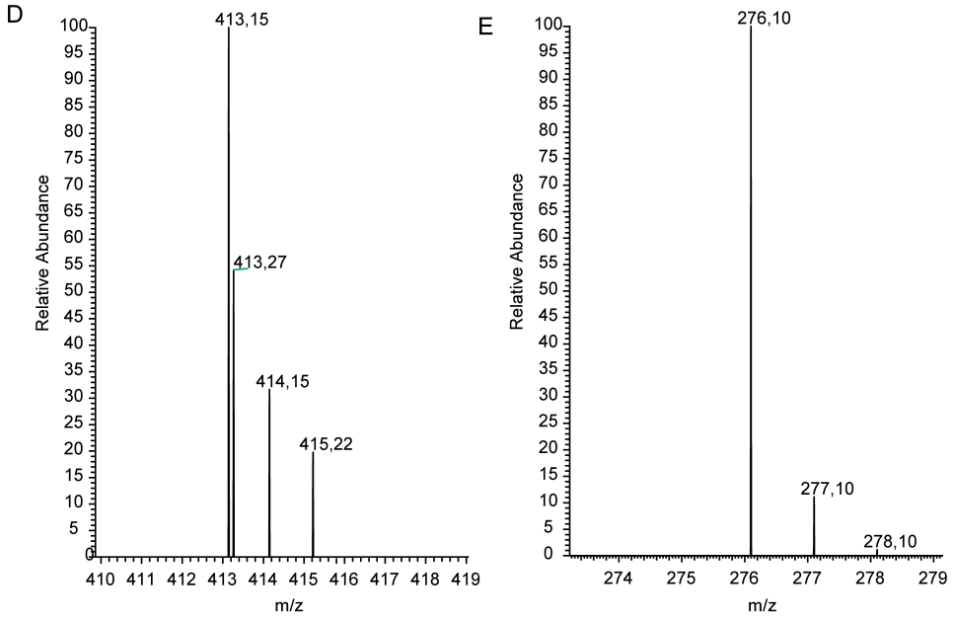


Figure S4. Mass spectra of the uncharacterized chrysogine related compounds. Compound **5** (A), compound **6** (B), compound **7** (C), compound **9** and **10** (D), compound **12** (E) – continued from page 75.

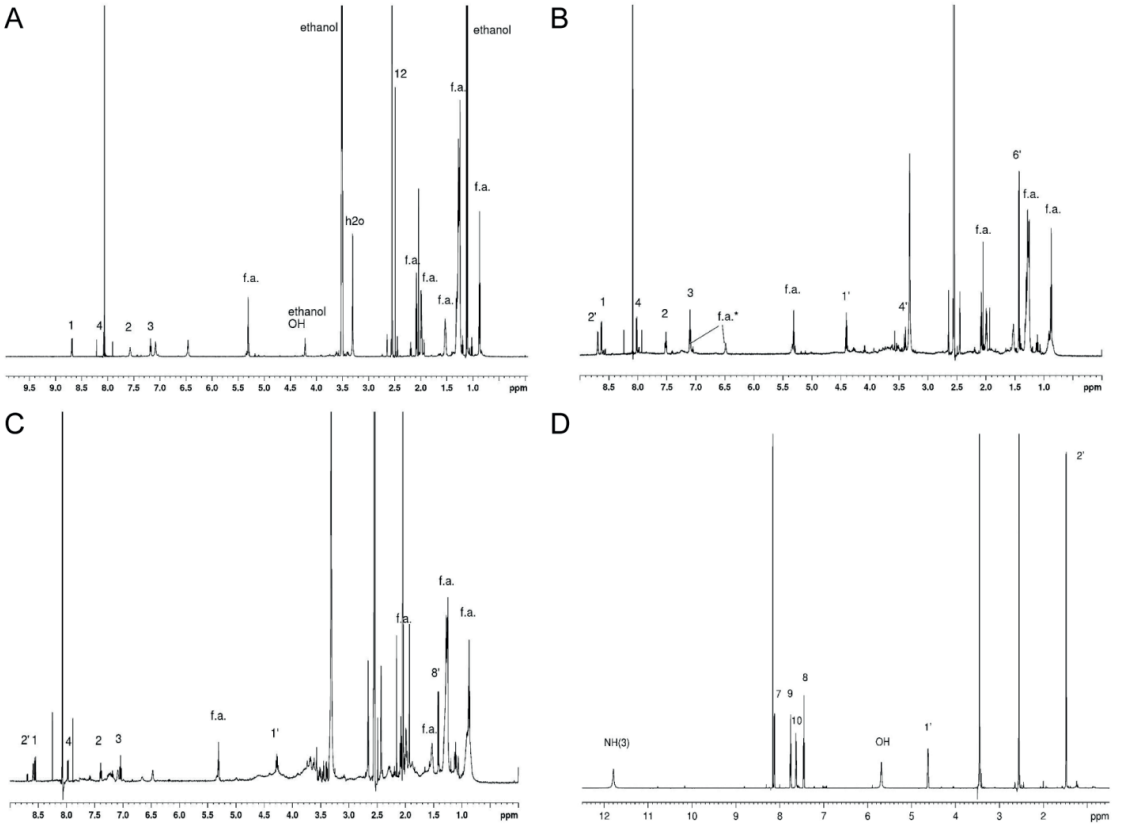


Figure S5. ^{1}H NMR spectrum of compound **3** (A), compound **8** (B), compound **13** (C), compound **1** (D), compound **4** (E). Compound **2** was observed as a minor impurity in this fraction. Signals are labelled with *– continues on page 78.

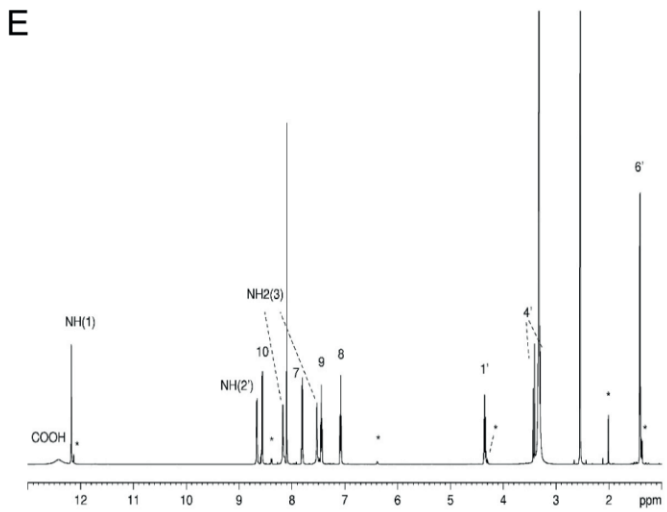


Figure S5. ^{1}H NMR spectrum of compound **3** (A), compound **8** (B), compound **13** (C), compound **1** (D), compound **4** (E). Compound **2** was observed as a minor impurity in this fraction. Signals are labelled with *– continued from page 77.

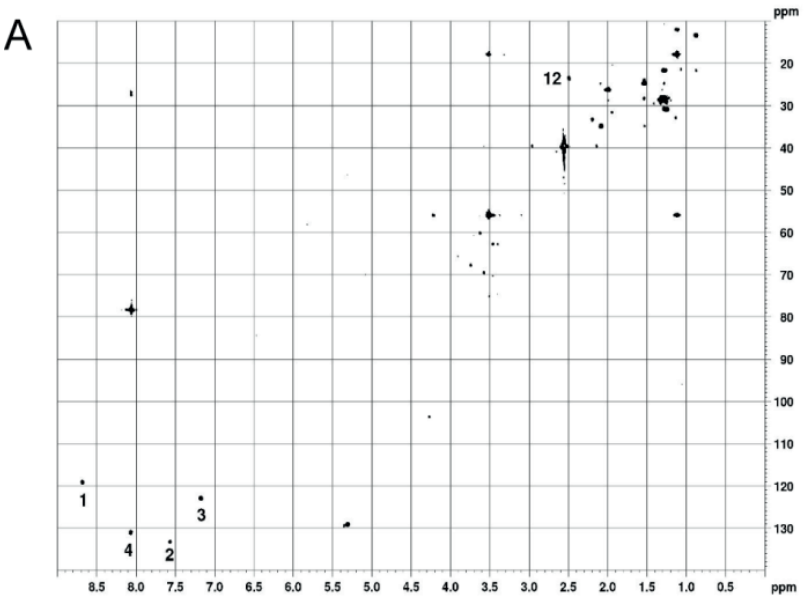


Figure S6. HSQC spectrum of compound **3** (A), compound **8** (B) and **13** (C) – continues on page 79.

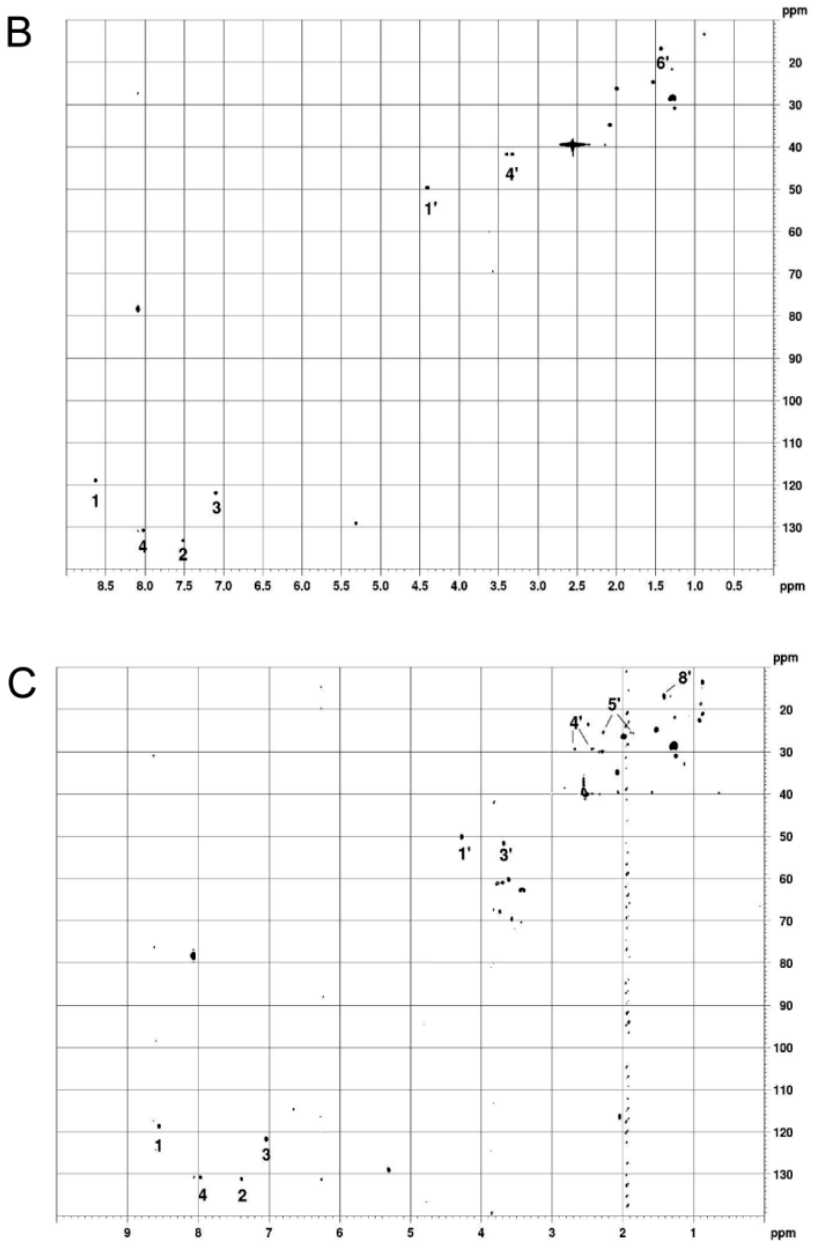


Figure S6. HSQC spectrum of compound **3** (A), compound **8** (B) and **13** (C) – continued from page 78.

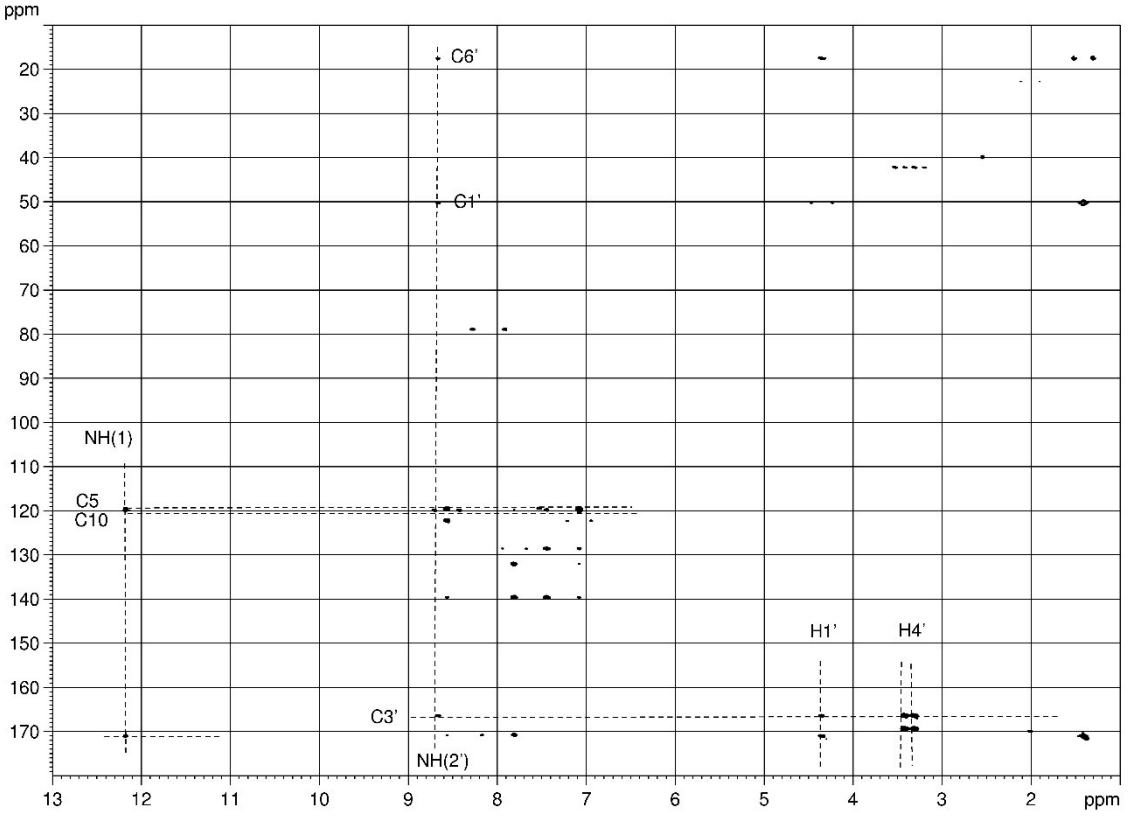
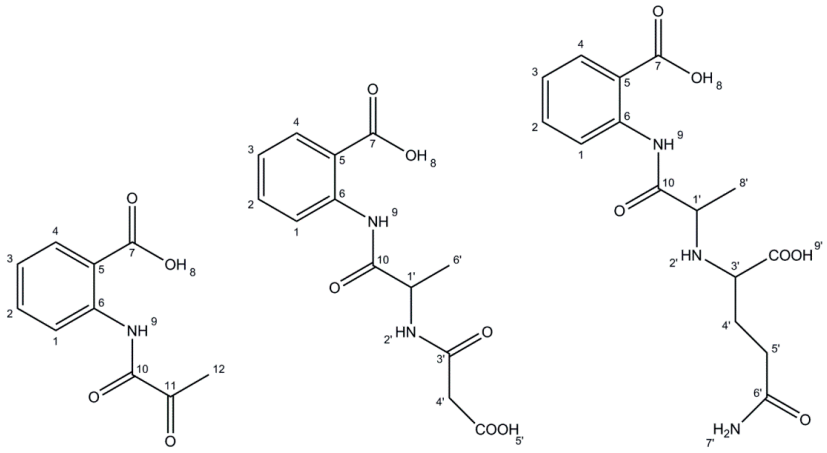


Figure S7. HMBC spectrum of compound 4 and impurity compound 2.

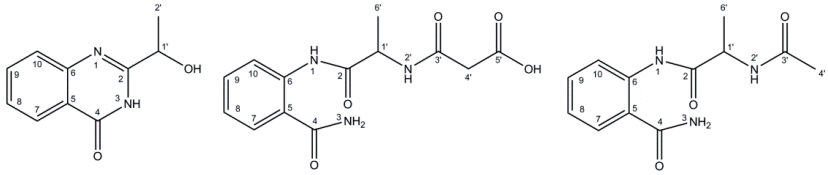


3 8 13

Table S1. Chemical shifts of compound **3**, compound **8** and compound **13** in DMSO/$CDCl_3$ 1/1. δ DMSO = 39.5 / 2.55 ppm. Temperature = 300 K.

	^{1}H	^{13}C	^{1}H	^{13}C	^{1}H	^{13}C
1	8.69	119.1	8.63	118.9	8.56	118.7
2	7.57	133.2	7.52	133.1	7.39	131.2
3	7.18	122.9	7.10	122.0	7.04	121.7
4	8.07	131.0	8.02	130.7	7.97	130.8
5	-	119.5	-	119.4	-	n.o.
6	-	139.1	-	140.4	-	138.8
7	-	168.9	-	169.3	-	168.9
8	n.o.	-	n.o.	-	n.o.	-
9	12.52	-	12.01	-	12.05	-
10	-	158.1	-	170.9	-	171.1
11	-	195.6				
12	2.49	23.6				
1’			4.41	49.8	4.28	50.2
2’			8.69	-	8.59	-
3’			-	166.2	3.69 (*)	51.7
4’			3.39 / 3.33	41.8	2.28 / 1.88 (*)	25.4
5’			-	169.0	2.68 / 2.44 (*)	29.4
6’			1.43	16.8	-	170.9
7’					n.o.	-
8’					1.42	16.9
9’					-	171.8

(*): multiplicities not resolved, -: not applicable, n.o.: not observed



1 4 2

Table S2. Chemical shifts of compound **1**, compound 4 and compound **2** in DMSO/$CDCl_3$ 1/1. δ DMSO = 39.5 / 2.55 ppm. Temperature = 280 K and 300 K.

	280 K	280 K	300 K		300 K			300 K	
	^{1}H	^{13}C	^{15}N	^{1}H	^{13}C	^{15}N	^{1}H	^{13}C	^{15}N
1	-	-	231.2	12.18	-	120.5	12.13	-	n.o.
2	-	159.7	-	-	171.0	-	-	n.o.	-
3	11.79	-	156.6	8.17 / 7.53	-	108.4	8.15 / 7.48	-	n.o.
4	-	161.7	-	-	170.8	-	-	n.o.	-
5	-	121.3	-	-	119.4	-	-	n.o.	-
6	-	148.5	-	-	139.6	-	-	n.o.	-
7	8.13	125.8	-	7.81	128.2	-	7.80	n.o.	-
8	7.46	126.0	-	7.08	121.9	-	7.07	n.o.	-
9	7.76	134.0	-	7.44	131.6	-	7.43	n.o.	-
10	7.63	126.9	-	8.56	119.5	-	8.59	n.o.	-
1’	4.63	67.2	-	4.35	49.7	-	4.31	n.o.	-
2’	1.48	21.8	-	8.67	-	125.7	8.39	-	125.3
OH	5.69	-	-						
3’				-	166.4	-	-	n.o.	-
4’				3.42 / 3.08	41.8	-	2.01	22.2	-
5’				-	169.3	-			
6’				1.41	17.0	-	1.38	n.o.	-
5’ COOH				12.42	-	-			

-: not applicable, n.o.: not observed

REFERENCES

1. Frisvad JC, Smedsgaard J, Larsen TO, Samson RA. 2004. Mycotoxins , drugs and other extrolites produced by species in *Penicillium* subgenus *Penicillium*. Stud Mycol 49:201–241.
2. Frisvad JC, Samson RA. 2004. Polyphasic taxonomy of *Penicillium subgenus Penicillium* A guide to identification of food and air-borne terverticillate *Penicillia* and their mycotoxins. Stud Mycol 49:1–174.
3. Pagano MC, Dhar PP. 2015. Fungal pigments: An overview, p. 173–181. *In* Fungal Biomolecules: Sources, Applications and Recent Developments.
4. Yunianto P, Rusman Y, Saepudin E, Priyono W, Sumaryono W. 2014. Alkaloid (Meleagrine and Chrysogine) from endophytic fungi (*Penicillium* sp.) of *Annona squamosa* L. Pak J Biol Sci 17:667–674.
5. Hikino H, Nabetani S, Takemoto T. 1973. Structure and biosynthesis of chrysogine, a metabolite of *Penicillium chrysogenum*. Yakugaku Zasshi 93:619–623.
6. Kimura Y, Inoue T, Tamura S. 1973. Isolation of 2-pyruvoylaminobenzamide as an antiauxin from *Colletotrichum lagenarium*. Agric Biol Chem 37:2213–2214.
7. Salo O, Ries M, Medema MH, Lankhorst PP, Vreeken RJ, Bovenberg RAL, Driessen AJM. 2015. Genomic mutational analysis of the impact of the classical strain improvement program on β–lactam producing *Penicillium chrysogenum*. BMC Genomics 16.
8. Pohl C, Kiel JAKW, Driessen AJM, Bovenberg RAL, Nygård Y. 2016. CRISPR/Cas9 Based Genome Editing of *Penicillium chrysogenum*. ACS Synth Biol 5:754–764.
9. Wollenberg RD, Saei W, Westphal KR, Klitgaard CS, Nielsen LK, Lysøe E, Gardiner DM, Wimmer R, Sondergaard TE, Sørensen JL. 2017. Chrysogine Biosynthesis Is Mediated by a Two-Module Nonribosomal Peptide Synthetase. J Nat Prod 6–10.
10. Keller NP, Turner G, Bennett JW. 2005. Fungal secondary metabolism - from biochemistry to genomics. Nat Rev Microbiol 3:937–947.
11. van den Berg M, Albang R, Albermann K, Badger J, Daran J, Driessen A, Garcia-Estrada C, Fedorova N, Harris D, Heijne W, Joardar V, Kiel J, Kovalchuk A, Martín J, Nierman W, Nijland J, Pronk J, Roubos J, van der Klei I, van Peij N, Veenhuis M, von Döhren H, Wagner C, Wortman J, Bovenberg R. 2008. Genome sequencing and analysis of the filamentous fungus *Penicillium chrysogenum*. Nat Biotechnol 26:1161–1168.
12. Guti S, Fierro F, Casqueiro J, Mart JF. 1999. Gene organization and plasticity of the β -lactam genes in different filamentous fungi. Antonie Van Leeuwenhoek 81–94.
13. Ali H, Ries MI, Nijland JG, Lankhorst PP, Hankemeier T, Bovenberg RAL, Vreeken RJ, Driessen AJM. 2013. A Branched Biosynthetic Pathway Is Involved in Production of Roquefortine and Related Compounds in *Penicillium chrysogenum.* PLoS One 8.
14. Ali H, Ries MI, Lankhorst PP, Hoeven RAM Van Der, Schouten OL, Peij NME Van, Bovenberg RAL, Vreeken RJ, Driessen AJM. 2014. A Non-Canonical NRPS Is Involved in the Synthesis of Fungisporin and Related Hydrophobic Cyclic Tetrapeptides in *Penicillium chrysogenum*. PLoS One 9.
15. Hoff B, Kamerewerd J, Sigl C, Zadra I, Kück U. 2010. Homologous recombination in the antibiotic producer *Penicillium chrysogenum* : strain Δ*Pcku70* shows up-regulation of genes from the HOG pathway. Appl Microbiol Biotechnol 85:1081–1094.
16. Snoek ISI, Krogt ZA Van Der, Touw H, Kerkman R, Pronk JT, Bovenberg RAL, Berg

MA Van Den, Daran JM. 2009. Construction of an *hdfA Penicillium chrysogenum* strain impaired in non-homologous end-joining and analysis of its potential for functional analysis studies. Fungal Genet Biol 46:418–426.

17. Kovalchuk A, Weber SS, Nijland JG, Bovenberg RAL, Driessen AJM. 2012. Chapter 1 Fungal ABC Transporter Deletion and Localization Analysis. Methods Mol Biol 835:1–16.
18. Kolar M, Puntb PJ, van den Hondel CAMJJ, Schwab H. 1988. Transformation of *Penicillium chrysogenum* using dominant selection markers and expression of an *Escherichia coli lacZ* fusion gene. Gene 62:127–134.
19. Weber E, Engler C, Gruetzner R, Werner S, Marillonnet S. 2011. A Modular Cloning System for Standardized Assembly of Multigene Constructs. PLoS One 6.
20. Bovenberg RAL, Kiel JAKW, Wenzel TJ, Los AP. 2014. US 2014/0106398 A1.
21. Gems D, Johnstone IL, Clutterhuck AJ. 1991. An autonomously replicating plasmid transforms *Aspergillus nidulans* at high frequency. Gene 98:61–67.
22. Onuki H, Miyashige H, Hasegawa H, Yamashita S. 1998. NI15501A, a Novel Anthranilamide Derivative from a Marine Fungus *Penicillium* sp. J Antibiot (Tokyo) 51:442–444.
23. Nielsen JC, Grijseels S, Prigent S, Ji B, Dainat J, Nielsen KF, Frisvad JC, Workman M, Nielsen J. 2017. Global analysis of biosynthetic gene clusters reveals vast potential of secondary metabolite production in *Penicillium species*. Nat Microbiol 17044.
24. Brakhage AA. 2012. Regulation of fungal secondary metabolism. Nat Rev Microbiol 11:21–32.
25. Bergmann S, Funk AN, Scherlach K, Schroeckh V, Shelest E, Horn U, Hertweck C, Brakhage AA. 2010. Activation of a silent fungal polyketide biosynthesis pathway through regulatory cross talk with a cryptic nonribosomal peptide synthetase gene cluster. Appl Environ Microbiol 76:8143–8149.
26. Shwab EK, Bok JW, Tribus M, Galehr J, Graessle S KN. 2007. Histone Deacetylase Activity Regulates Chemical Diversity in *Aspergillus*. Eukaryot Cell 6:1656–1664.
27. Weber SS, Kovalchuk A, Bovenberg RAL, Driessen AJM. 2012. The ABC transporter ABC40 encodes a phenylacetic acid export system in *Penicillium chrysogenum*. Fungal Genet Biol 49:915–921.
28. Harju S, Fedosyuk H, Peterson KR. 2004. Rapid isolation of yeast genomic DNA : Bust n' Grab. BMC Biotechnol 6:1–6.
29. Salo O, Guzmán-Chávez F, Ries MI, Lankhorst PP, Bovenberg RAL, Vreeken RJ, Driessen AJM. 2016. Identification of a polyketide synthase involved in sorbicillin biosynthesis by *Penicillium chrysogenum*. Appl Environ Microbiol 82:3971–3978.